UPDATING AMBIGUOUS BELIEFS[*]

by

Itzhak Gilboa[**]

and

David Schmeidler[***]

Extended Abstract

November 1991

[**] Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University, Evanston, Illinois 60208.

[***] Department of Economics, Ohio State University, and Department of Statistics and Department of Economics, Tel Aviv University, Tel Aviv, Israel.

The Bayesian approach to decision making under uncertainty prescribes that a decision maker have a unique prior probability and a utility function such that decisions are made so as to maximize the expected utility. In particular, in a statistical inference problem the decision maker is assumed to have a probability distribution over all possible distributions which may govern a certain random process.

This paradigm was justified by axiomatic treatments, most notably that of Savage (1954), and it enjoys unrivaled popularity in economic theory, game theory, and so forth.

However, this theory is challenged by two classes of evidence: on the one hand, there are experiments and mind experiments (such as Ellsberg (1961) and many others) which seem to show that individuals tend to violate the consistency conditions underlying (and implied by) the Bayesian approach. On the other hand, people seem to have difficulties with specification of a prior for actual statistical inference problems. Thus, classical—rather than Bayesian—methods are used for practical purposes, although they are theoretically less satisfactory.

The last decade has witnessed--among numerous and various general-izations of von Neumann-Morgenstern's (1947) expected utility theory--generalizations of the Bayesian paradigm as well. We will not attempt to provide a survey of them here. Instead, we only mention the models which are relevant to the sequel.

1.   <u>Non-additive Probabilities</u>. First introduced by Schmeidler (1982, 1986, 1989), and also axiomatized in Gilboa (1987). Fishburn (1988), and Wakker (1989), non-additive probabilities are monotone set-functions which

do not have to satisfy additivity. Using the Choquet integral (Choquet (1953-54)) one may define expected utility, and the works cited before axiomatize preference relations which are representable by expected utility in this sense.

2. <u>Multiple Priors</u>. As axiomatized by Gilboa and Schmeidler (1989), this model assumes that the decision maker has a <u>set</u> of priors, and each alternative is assessed according to its minimal expected utility, where the minimum is taken over all priors in the set. (This idea is also related to Bewley (1986, 1987, 1988), who suggests a partial order over alternatives, such that one alternative is preferred to another only if its expected utility is higher according to <u>all</u> priors in the set.)

Of particular interest to this study is the intersection of the two models: it turns out that if a non-additive measure exhibits uncertainty aversion (technically, if it is convex in the sense

$$v(A \cup B) + v(A \cap B) \geq v(A) + v(B)$$

then the Choquet integral of a real-valued function with respect to v equals the minimum of all its integrals with respect to additive priors taken from the core of v. (The core is defined as in cooperative game theory, i.e., p is in the core of v is $p(A) \geq v(A)$ for every event A with equality for the whole sample space. Convex non-additive measures have nonempty cores.)

While these models--as many others--suggested generalizations of the Bayesian approach for a one-shot decision problem, they shed very little light on the problem of dynamically updating probabilities as new information is gathered. We find this problem to be of paramount importance

for several interrelated reasons:

1.   The theoretical validity of any model of decision making under uncertainty is quite dubious if it cannot cope successfully with the dynamic aspect.

2.   The updating problem is at the heart of statistical theory.   In fact, it may be viewed as the problem statistical inference is trying to solve.   Some of the works in the statistical literature which pertain to this study are Agnew (1985), Genest and Schervish (1985), and Lindley, Tversky and Brown (1979).

3.   Applications of these models to economic and game theory models require some assumptions on how economic agents change their beliefs over time.   The question naturally arises, then: What are reasonable ways to update such beliefs?

4.   The theory of artificial intelligence, which in general seems to have much in common with the foundations of economic, decision and game theory, also tries to cope with this problem.   See, for instance, Fagin and Halpern (1989), Halpern and Fagin (1989), and Halpern and Tuttle (1989).

In this study we try to deal with the problem axiomatically and suggest plausible update rules which satisfy some basic requirements.   We present a family of pseudo-Bayesian rules, each of which may be considered a generalization of Bayes' rule for a unique additive prior.   We characterize those that map the set of non-additive preferences onto itself, and two special cases interpreted as "optimistic" and "pessimistic" updates single themselves out.   We also present a family of "classical" update rules, each of which starts out with a given set of priors, possibly rules some of them

out in the face of new information, and continues with the (Bayesian) updates of the remaining ones.

In particular, a maximum-likelihood update rule would be the following: consider only those priors which ascribe the maximal probability to the event that is known to have occurred, update each of them according to Bayes' rule and continue in this fashion.

It turns out that if the set of priors one starts out with can also be represented by a non-additive probability, the results of this rule are independent of the order in which information is gathered.

Furthermore, for those preferences which can be simultaneously represented by a non-additive measure and by multiple priors, the maximum likelihood update rule coincides with one of the more intuitive Bayesian rules--namely, the "pessimistic one"--and they boil down to the Dempster-Shafer rule (see Dempster (1967, 1968), Shafer (1976), and Smets (1986)). For recent work on belief functions (which are, in particular, convex) and their updating, see Jaffray (1989), Chateauneuf and Jaffray (1989), and especially Jaffray (1990).

Thus, we find that an axiomatically-based generalization of the Bayesian approach can accommodate multiple priors (which are used in classical statistics). Moreover, the maximum likelihood principle, which is at the heart of statistical inference (and implicit in the techniques of confidence sets and hypotheses testing) coincides with the generalized Bayesian updating.

Due to the prominence of this rule, it may be insightful to study it in a simple example. Consider Ellsberg's example in which an urn with 90 balls is given, out of which 30 are red, and 60 are either blue or yellow. For

simplicity of exposition, let us model this situation in a somewhat extreme fashion, allowing for <u>all</u> distributions of blue and yellow balls. Maxmin expected utility with respect to this set of priors is equivalent to the maximization of the Choquet integral of utility w.r.t. to a non-additive measure v defined as follows:

$$v(R) = 1/3 \qquad\qquad v(B) = v(Y) = 0$$
$$v(R \cup B) = v(R \cup Y) = 1/3 \qquad v(B \cup Y) = 2/3$$
$$v(R \cup B \cup Y) = 1$$

where R, B and Y denote the events of a red, blue or yellow ball being drawn, respectively.

Assume know that it is known that a ball (which, say, has already been drawn) is not red. Conditioning on the event B ∪ Y, <u>all</u> priors in the set ascribe probablity of 2/3 to it. Thus, they are all left in the set and updated according to Bayes' rule. This captures our intuition that no ambiguity was resolved, and our complete ignorance regarding the event B ∪ Y has not changed. (Actually, it is now highlighted by the fact that this event, about which we know the least, is now known to have occurred.)

Consider, on the other hand, the same update rule in the case that R ∪ B is known. The priors we started out with ascribe to this event probablities ranging from 1/3 to 2/3. According to the maximum likelihood principle, only one of them is chosen--namely, the p which satisfies

$$p(R) = 1/3 \qquad\qquad p(B) = 2/3 \qquad\qquad p(Y) = 0.$$

In this particular case, the set of priors shrinks to a singleton and, equivalently, the updated measure v is additive (and equals p itself). Ambiguity is thus reduced (in this case, eliminated) by the generalized Bayesian learning embodied in the exclusion of some priors.

In the context of such examples it is sometimes argued that the maximum likelihood rule is too extreme, and that priors which, say, only ε-maximize the likelihood function should not be ruled out. Indeed, classical statistics techniques such as hypotheses testing do allow for ranges of the likelihood function.

At present we are not aware of a nice axiomatization of such rules. We would like to point out, however, that the other extreme rule--i.e., updating all priors without excluding any of them (see, for instance, Fagin and Halpern (1989), and Jaffray (1990))--does not appear to be any less "extreme" in general, nor does it seem to be implied by more compelling axioms.

To illustrate the two extreme approaches, one may think of maxmin expected utility with a non-unique prior as describing a behavior of a cautious decision maker who has several consultants. Each of these has a (unique, additive) prior, and the decision maker evaluates a decision by a minimal expected utility assigned to it by the various consultants. The "Full Bayesian Update" (Fagin and Halpern (1989), Jaffray (1990)) suggests that no matter how many times a certain consultant was proven to be "wrong" in the sense of assigning (relatively) low probabilities to events that actually occurred, this consultant's advice will still be heeded. The Dempster-Shafer rule we axiomatize here suggests that once a consultant fails to be among the maximizers of the probability assessment of an event

that has occurred, he/she is instantly fired. Thus stated, both approaches seem, indeed, rather extreme.

Let us emphasize, however, that the consultant's story should only be taken as a metaphor. Since the update rule is derived from axioms regarding behavior, a decision maker may not be aware of these priors at all--indeed, not even of his/her own utility. Reasonability as well as desirability of various decision rules is typically more lucidly examined in view of the more "primitive" axioms characterizing them than in isolation. From this viewpoint, the "pessimistic" axiom that gives rise to our update rule seems to be in agreement with the "pessimistic" axioms that gave rise to the multiple-prior model in the first place (see Gilboa-Schmeidler (1989)).

We believe that our theory can be applied to a variety of economic models, explaining phenomena which are incompatible with the Bayesian theory, and possibly providing better predictions. As a matter of fact, this belief may be updated given new evidence: Dow and Werlang (1990), and Simonsen and Werlang (1990) have already applied the multiple prior theory to portfolio selection problems. These studies have shown that a decision maker having ambiguous beliefs will have a (non-trivial) range of prices at which he/she will neither buy nor sell an uncertain asset, exhibiting inertia in portfolio selection. Applying our new results regarding ambiguous beliefs update, one may study the conditions under which these price ranges will shrink in the face of new information.

Dow, Madrigal and Werlang (1989) studied trade among agents, at least one of whom has ambiguous beliefs. They show that the celebrated no-trade result of Milgrom and Stokey (1982) fails to hold in this context. In this study, the Dempster-Shafer rule for updating non-additive measures was used,

a rule which is justified by the current paper. Casting the trade set-up

into a dynamic context raises the question of an asymptotic no-trade

theorem: Under what conditions will additional information reduce the

volume or probability of trade?

In another recent study, Yoo (1990) addressed the question of why stock

prices tend to fall after the initial public offering and rise at a later

stage. Although Yoo uses ambiguous beliefs mostly as in Bewley's (1986)

model, his results can also be obtained using the models mentioned above.

It seems that the update rule justified by our study may explain the price

dynamics.

These various models seem to point at a basic problem: given a convex

non-additive measure (or, equivalently, a set of priors which is the core of

such a measure), under what conditions will the Dempster-Shafer rule yield

convergence of beliefs to a single additive prior? Obviously, the answer

cannot be "always." Consider a "large" measurable space with all possible

priors (equivalently, with the "unanimity game" as a non-additive measure).

In this set-up of "complete ignorance," no conclusions about the future may

be drawn from past observations—that is, the updated beliefs still include

all possible priors. However, with some initial information (say, finitely

many extreme points of the set of priors) convergence is possible. We

intend to study conditions that will guarantee convergence in future works.

2.   Framework and Preliminaries

Let X be a set of consequences endowed with a weak order $\geq$. Let (S,Σ)

be a measurable space of states of the world, where Σ is the algebra of

events. A function f: S $\rightarrow$ X is Σ-measurable if for every x $\in$ X

$$\{s|f(s) > x\}, \{s|f(s) \geq x\} \in \Sigma.$$

Let $F = \{f: S \to X|f \text{ is } \Sigma\text{-measurable}\}$ be the set of <u>acts</u>.

Let $F_0 = \{f \in F| \ |range(f)|< \infty\}$ be the set of <u>simple acts</u>.

A function $u: X \to \mathbb{R}$, which represents $\geq$, i.e.,

$$u(x) \geq u(y) \iff x \geq y, \ \forall \ x,y \in X$$

is called a <u>utility</u> function.

A function $v: \Sigma \to [0,1]$ satisfying

(i) $v(\emptyset) = 0; v(S) = 1;$

(ii) $A \subseteq B \Rightarrow v(A) \leq v(B)$

is a <u>non-additive measure</u>. It is <u>convex</u> if

$$v(A \cup B) + v(A \cap B) \geq v(A) + v(B)$$

for all $A,B \in \Sigma$. It is <u>additive</u>, or simply a <u>measure</u> if the above inequality is always satisfied as an equality.

A real valued function is Σ-measurable if for every $t \in \mathbb{R}$

$$\{s|w(s) \geq t\}, \{s|w(s) > t\} \in \Sigma.$$

Given such a function w and a non-additive measure v, the (Choquet) integral of w w.r.t. (with respect to) v on S is

$$\int w dv = \int_S w dv = \int_0^\infty v(\{s \mid w(s) \geq t\}) \, dt + \int_{-\infty}^0 [v(\{s \mid w(s) \geq t\}) - 1] dt.$$

For a non-additive measure v we define the core as for a cooperative game, i.e.,

$$Core(v) = \{p \mid p \text{ is a measure s.t. } p(A) \geq v(A) \ \forall \ A \in \Sigma\}.$$

Recall that a convex v has a nonempty core (see Shapley (1965)).

We are now about to define two classes of binary relations on F: those represented by maximization of expected utility with non-additive measures (NA), and those represented by maxmin of expected utility with multiple priors (MP).

Denote by NA_o ($= NA_o(X, \geq, S, \Sigma)$) the set of binary relations $\geq$ on F such that there are a utility u, unique up to p.l.t. (positive linear transformation), and a unique non-additive measure v satisfying:

(i)   for every $f \in F$, u o f is Σ-measurable;

(ii)   for every f, g, $\in$ F.

$f \geq g \iff \int u \ o \ f \ dv \geq \int u \ o \ g \ dv.$

Note that in general the measurability of f does not guarantee that of u o f, and that (ii) implies that $\geq$ on F, when restricted to constant functions, extends $\geq$ on X. Hence we use this convenient abuse of notation. Similarly, we will not distinguish between $x \in X$ and the constant act which equals x on S.

Characterizations of NA_o were given by Schmeidler (1986, 1989) for the Anscombe-Aumann (1963) framework, where X is a mixture space and u is assumed affine, by Gilboa (1987) in the Savage (1954) framework, where X is

arbitrary but $\Sigma = 2^S$ and v is nonatomic, and by Wakker (1989) for the case where X is a connected topological space. Fishburn (1988) extended the characterization to non-transitive relations.

Let MP_o (= $MP_o(X,\geq,S,\Sigma)$) denote the set of binary relations $\geq$ on F such that there are a utility u unique up to a p.l.t., and a unique nonemtpy, closed (in the weak* topology) and convex set C of (finitely additive) measures on Σ such that:

(i)   for every f $\in$ F, u o f is Σ-measurable;

(ii)   for every f,g $\in$ F

$$f \geq g \iff \min_{p \in C} \int u \circ f \, dp \geq \min_{p \in C} \int u \circ g \, dp.$$

A characterization of MP_o in the Anscombe-Aumann framework was given in Gilboa-Schmeidler (1989). To the best of our knowledge, there is no such axiomatization in the framework of Savage. However, the set $NA_o \cap MP_o$, which will play an important role in the sequel may be characterized by strengthening the axioms in Gilboa (1987).

It will be convenient to include the trivial weak order $\geq^* = F \times F$ in both NA and MP. Hence, we define $NA = NA_o \cup \{\geq^*\}$ and $MP = MP_o \cup \{\geq^*\}$.

For simplicity we will assume that X has $\geq$-maximal and minimal elements. More specifically, let $x^*, x_* \in X$ satisfy $x_* \leq x \leq x^*$ for all $x \in X$. W.l.o.g. (without loss of generality), assume that x_* and x^* are unique. Since for both NA_o and MP_o the utility function is unique up to a p.l.t. we will assume w.l.o.g. that $u(x_*) = 0$ and $u(x^*) = 1$ for all utilities henceforth considered.

When X is a mixture space we define NA' and MP' to be the subsets of NA

and MP, respectively, where the utility u is also required to be affine.
For such spaces X we recall the following results.

Proposition 2.1: Suppose that $\geq$ $\in$ NA' and let v be the associated non-additive measure. Then $\geq$ $\in$ MP' iff v is convex.

Proposition 2.2: Suppose that $\geq$ $\in$ MP' and let C be the associated set of measures. Define

$$v(A) = \min_{p \in C} p(A) \text{ for } A \in \Sigma.$$

Then v is a non-additive measure and $\geq$ $\in$ NA' iff v is convex and
C = Core(v).

The proofs of these appear, explicitly or implicitly, in Schmeidler
(1984, 1986, 1989). Note that the axiomatization of NA' (Schmeidler (1989))
uses comonotomic independence, and given this property the convexity of v is
equivalent to uncertainty aversion. The axiomatization of MP' (Gilboa and
Schmeidler (1989)) uses a weaker independence notion--termed C-independence-
-and uncertainty aversion. Given these, the convexity of v and the equality
C = Core(v) (where v is defined as in Proposition 2.2) is equivalent to
comonotomic independence.

We now turn to define update rules. We need the following definitions.
Given a measurable partition $\Pi = \{A_i\}_{i=1}^{n}$ of S and $\{f_i\}_{i=1}^{n} \subseteq F$, let
$(f_1,A_1;\ldots;f_n,A_n)$ denote the act g $\in$ F satisfying $g(s) = f_i(s)$ for all
s $\in A_i$ and all $1 \leq i \leq n$. Given a binary relation $\geq$ on F, an event A $\in \Sigma$ is

$\geq$-null iff the following holds: for every $f, g, h_1, h_2 \in F$,

$$f \geq g \text{ iff } (f, A^C; h_1, A) \geq (g, A^C; h_2, A).$$

Let $\bar{\mathcal{B}}$ denote the set of all binary relations on F. Given $\mathcal{B} \subseteq \bar{\mathcal{B}}$, an update rule for $\mathcal{B}$ is a collection of functions, $U = \{U_A\}_{A \in \Sigma}$, where $U_A: \mathcal{B} \to \bar{\mathcal{B}}$ such that for all $\geq \in \mathcal{B}$ and $A \in \Sigma$, A^C is $U_A(\geq)$ - null and $U_S(\geq) = \geq$. $U_A(\geq)$ should be thought of as the preference relation once A is known to have occurred.

Given $\mathcal{B}$ and an update rule for it, $U = \{U_A\}_{A \in \Sigma}$, U is said to be commutative w.r.t. $\geq$ or $\geq$-commutative if for every $A, B \in \Sigma$ we have

$$U_B(U_A(\geq)) = U_{A \cap B}(\geq).$$

It is commutative if it is commutative w.r.t. $\geq$ for all $\geq \in \mathcal{B}$. (Note that this condition is stronger than strict commutativity, i.e., $U_A \circ U_B = U_B \circ U_A$. However, "commutativity" seems to be a suggestive name which is not overburdened with other meanings.)

3.   Bayesian and Classical Rules

Given a set $\mathcal{B}$ of binary relations on F, every $f \in F$ suggests a natural update rule for $\mathcal{B}$:   define $BU^f = \{BU_A^f\}_{A \in \Sigma}$ by

$$g \; BU_A^f(\geq) \; h \; <=> \; (g, A; f, A^C) \geq (h, A; f, A^C) \text{ for all } g, h \in F.$$

It is obvious that for every f, BU^f is an update rule, i.e., that A^C is

$BU^f_A(\geq)$ - null for all $\geq \in \mathcal{B}$ and $A \in \Sigma$. We will refer to it as the f-Bayesian update rule and $\{BU^f\}_{f\in F}$ will be called the set of Bayesian update rules.

Note that for $\geq \in NA$ with an additive v, all the Bayesian update rules coincide with Bayes' rule, hence the definition of the Bayesian update rules may be considered a formulation and axiomatization of Bayes' rule in the case of (a unique) additive prior.

Proposition 3.1: For every $\geq \in \mathcal{B}$ and $f \in F$, BU^f is $\geq$-commutative.

Theorem 3.2: Let $f \in F$ and assume that $|\Sigma| > 4$. Then the following are equivalent:

    (i) $BU^f_A(NA') \subseteq NA'$ for all $A \in \Sigma$;

    (ii) $f = (x^*,T;x_*,T^C)$ for some $T \in \Sigma$.

Of particular interst are the Bayesian update rules corresponding to $f = x^*$ and $f = x_*$(i.e., $T = S$ or $T = \emptyset$ in (ii) above). For the latter (x_*) there is an "optimistic" interpretation: when comparing two actions given a certain event A, the decision maker implicitly assumes that had A not occurred, the worst possible outcome (x_*) would have resulted. In other words, the behavior given A--$BU^f_A(\geq)$--exhibits "happiness" that A has occurred; the decisions are made as if we are always in "the best of all possible worlds."

Note that the corresponding non-additive measure is

$$v_A(B) = v(B \cap A)/v(A).$$

On the other hand, for f = x*, we consider a "pessimistic" decision

maker, whose choices reveal the hidden assumption that all the impossible

worlds are the best possible ones. This rule defines the non-additive

function by

$$v_A(B) = [v((B \cap A) \cup A^c) - v(A^c)] / (1 - v(A^c)),$$

which is identical to the Dempster-Shafer rule for updating probabilities.

It should not surprise us that this "pessimistic" rule is going to play

a major role in relation to MP--i.e., to uncertainty-averse decision makers

who follow a maxmin (expected utility) decision rule. In a similar way one

may develop a "dual" theory of "optimism" in which uncertainty-seeking will

replace uncertainty-aversion, concavity of v will replace convexity, and

maxmax will supercede maxmin. For this "dual" theory, the update rule

$$v_A(B) = v(B \cap A)/v(A)$$

would be the "appropriate" one (in a sense that will be clear shortly).

Notice that this rule was used--without axiomatization--as a definition of

probability update in Gilboa (1989).

Taking a classical statistics' point of view, it is natural to start

out with a set of priors. Hence we only define classical update rules for

$\mathcal{B}$ = MP'. A natural procedure in the classical updating process is to rule

out some of the given priors, and update the rest according to Bayes' rule.

Thus, we get a family of update rules, which differ in the way the priors

are selected.

Formally, a <u>classical update rule</u> is characterized by a function

R: (C,A) $\mapsto$ C' such that C' $\subseteq$ C is a closed and convex set of measures for

every such C and every A $\in$ Σ, with R(C,S) = C. The associated update rule

will be denoted $CU^R = \{CU_A^R\}_{A \in \Sigma}$.   (If R(C,A) = $\emptyset$ we define $CU_A^R(\geq)$ = $\geq^*$.)

Note that these are indeed update rules, i.e., that for every $\geq$ $\in$ MP', every

R and every A $\in$ Σ, A^C is $CU_A^R(\geq)$ - null. Furthermore, for $\geq$ $\in$ MP' with an

associated set C, $CU_A^R(\geq)$ $\in$ MP' provided that inf$\{p(A)|p \in R(C,A)\}$ > 0 for

all A $\in$ Σ.

Of particular interest will be the following classical update rule

called maximum likelihood and defined by

$$R^0(C,A) = \{p \in C | p(A) = \max_{q \in C} q(A) > 0\}.$$

<u>Theorem 3.3</u>: CU^{R^0} is cummutative on NA' $\cap$ MP'. Furthermore, for

$\geq$ $\in$ NA' $\cap$ MP',

$$BU_A^{(x^*,S)}(\geq) = CU_A^{R^0}(\geq) \in NA' \cap MP'.$$

I.e., the Bayesian update rule with f = (x*,S) coincides with the maximum

likelihood classical update rule. Furthermore, they are also equivalent to

the Dempster-Shafer update rule for belief functions. (Note that every

belief function (see Shafer (1976)) is convex, though the converse is false.

Yet one may apply the Dempster-Shafer rule for every non-additive measure

v.)

References

Agnew, C. E., "Multiple Probability Assessments by Dependent Experts,"
Journal of the American Statistical Association, 80 (1985), 343-347.

Anscombe, F. J. and R. J. Aumann, "A Definition of Subjective Probability,"
The Annals of Mathematics and Statistics, 34 (1963), 199-205.

Bewley, T., "Knightian Decision Theory: Part I," Cowles Foundation
Discussion Paper No. 807, Yale University (1986).

Bewley, T., "Knightian Decision Theory: Part II: Intertemporal Problems,"
Cowles Foundation Discussion Paper No. 835, Yale University (1987).

Bewley, T., "Knightian Decision Theory and Econometric Inference," Cowles
Foundation Discussion Paper No. 868, Yale University (1988).

Chateauneuf, A. and J.-Y. Jaffray, "Some Characterizations of Lower
Probabilities and Other Monotone Capacities Through the Use of Moebius
Inversion," Mathematical Social Sciences, 17 (1989), 263-283.

Choquet, G., "Theory of Capacities," Annales de l'Institut Fourier, 5 (1953-
54), 131-295.

Dempster, A. P., "Upper and Lower Probabilities Induced by a Multivalued
Mapping," Annals of Mathematical Statistics, 38 (1967), 325-339.

Dempster, A. P., "A Generalization of Bayesian Inference," Journal of the
Royal Statistical Society, Series B 30 (1968), 205-247.

Dow, J. and S. R. C. Werlang, "Risk Aversion, Uncertainty Aversion and the
Optimal Choice of Portfolio," London Business School Working Paper
(1990).

Dow, J., V. Madrigal and S. R. C. Werlang, "Preferences, Common Knowledge
and Speculative Trade," Fundacao Getulio Vargas Working Paper (1989).

Ellsberg, D., "Risk, Ambiguity and the Savage Axioms," Quarterly Journal of Economics, 75 (1961), 643-669.

Fagin, R. and J. Y. Halpern, "A New Approach to Updating Beliefs," mimeo (1989).

Fishburn, P. C., "Uncertainty aversion and Separated Effects in Decision Making Under Uncertainty," in Combining Fuzzy Imprecision with Probabilistic Uncertainty in Decision Making, J. Kacprzyk and M. Fedrizzi (eds.), Springer-Verlag (1988), pp. 10-25.

Genest, C. and M. J. Schervish, "Modeling Expert Judgments for Bayesian Updating," The Annals of Statistics, 13 (1985), 1198-1212.

Gilboa, I., "Expected Utility Theory with Purely Subjective Non-Additive Probabilities," Journal of Mathematical Economics, 16 (1987), 65-88.

Gilboa, I., "Additivizations of Non-Additive Measures," Mathematics of Operations Research, 14 (1989), 1-17.

Gilboa, I. and D. Schmeidler, "Maxmin Expected Utility with Non-Unique Prior," Journal of Mathematical Economics, 18 (1989), 141-153.

Halpern, J. Y. and R. Fagin, "Two Views of Belief: Belief as Generalized Probability and Belief as Evidence," mimeo (1989).

Halpern, J. Y. and M. R. Tuttle, "Knowledge, Probability and Adversaries," mimeo (1989).

Jaffray, J.-Y., "Linear Utility Theory for Belief Functions," Operations Research Letters, 8 (1989), 107-112.

Jaffray, J.-Y., "Bayesian Updating and Belief Functions," mimeo (1990).

Lindley, D. V., A. Tversky and R. V. Brown, "On the Reconciliation of Probability Assessments," Journal of the Royal Statistical Society, Series A 142 (1979), 146-180.

Milgrom, P. and N. Stokey, "Information, Trade and Common Knowledge,"
Journal of Economic Theory, 26 (1982), 17-27.

von Neumann, J. and O. Morgenstern, Theory of Games and Economic Behavior,
2nd Ed., Princeton University Press, Princeton, N.J. (1947).

Savage, L. J., The Foundations of Statistics, Wiley, New York (1954).

Schmeidler, D., "Subjective Probability without Additivity" (temporary
title), working paper, Foerder Institute for Economic Research, Tel
Aviv University (1982).

Schmeidler, D., "Nonadditive Probabilities and Convex Games," mimeo (1984).

Schmeidler, D., "Integral Representation without Additivity," Proceedings of
the American Mathematical Society, 97 (1986), 253-261.

Schmeidler, D., "Subjective Probability and Expected Utility without
Additivity," Econometrica, 57 (1989), 571-587.

Shafer, G., A Mathematical Theory of Evidence, Princeton University Press,
Princeton, N.J. (1976).

Simonsen, M. H. and S. R. C. Werlang. "Subadditive Probabilities and
Portfolio Inertia," mimeo (1990).

Shapley, L. S., "Notes on n-person Games VII: Cores of Convex Games," The
Rand Corporation R. M. (1965); also as "Cores of Convex Games,"
International Journal of Game Theory, 1 (1971), 12-26.

Smets, Ph., "Combining Non-distinct Evidences," Technical Report
ULB-IRIDIA-86/3 (1986).

Wakker, P., "Continuous Subjective Expected Utility with Non-Additive
Probabilities," Journal of Mathematical Economics, 18 (1989), 1-27.

Yoo, K. R., "A Theory of the Underpricing of Initial Public Offerings,"
mimeo (1990).